Stable and Fast Techniques for Unambiguous Compound Phase Coding

Andrea Torsello, Andrea Albarelli, Emanuele Rodolà

Dipartimento di Scienze Ambientali, Informatica e Statistica

Università Ca' Foscari Venezia - via Torino, 155 - 30172 Venice Italy

Abstract

Phase shift methods have proven to be very robust and accurate for photometric 3D reconstruction. One problem of these approaches is the existence of ambiguities arising from the periodicity of the fringe patterns. While several techniques for disambiguation exist, all of them require the projection of a significant number of additional patterns. For instance, a global Gray coding sequence or supplemental sinusoidal patterns of different periods are commonly used to complement the basic phase shift technique. In this paper we propose four new coding strategies that encode the index of the projected column using several phases and that mix the resulting phases into a controllable number of projected patterns from which the position can be recovered with subpixel precision. One notable characteristic of the proposed approaches is that we can allocate the additional number of patterns specifically to improve precision or provide higher robustness to noise. The proposed approaches are analyzed and compared with the state of the art, showing their ability to be tuned towards high precision in low noise conditions or robustness with respect to noise.

1. Introduction

The main challenge for any triangulation-based surface reconstruction technique is the assignment of reliable correspondences between features observed by two or more different points of view. Given the central role of this problem, many and diverse strategies have been proposed in literature over the past

few decades [1]. When a sparse reconstruction is adequate, correspondences can be searched and tracked among repeatable features readily present in the scene, such as corners or edges. Unfortunately, in general it is not possible to guarantee that the same features are extracted from each image, or that the feature density is sufficient. Hence, complementary techniques, usually based on photometric correlation, are used to obtain an approximate reconstruction of the scene depth map. Other systems overcome these limitations as they do not rely on natural features, but instead use projected patterns of light in order to find correspondences that are usually as dense as the pixels of each image [2]. Such techniques can be much more accurate than feature-based approaches and have received particular attention from scientific and engineering communities in the fields of optics, metrology and computer science. Further, structured light approaches can be used almost unchanged in single camera systems, by calibrating the camera and the projector.

The main idea behind 3D reconstruction based on projected light is that of assigning unique codes to surface points by modulating the intensity of a light projected onto the object. These codes are used to extract correspondences between points in different views, which can then be triangulated to obtain depth estimates. To this end, several patterns are projected onto the scene in such a way that each point has a unique sequence of intensity values. These sequences are then decoded to obtain the point's identity. The main issue in structured light approaches is the design of projection schemes that allow for robust and precise decoding, possibly with subpixel accuracy, using the smallest number of patterns possible while still guaranteeing reliable measurements. Such an endeavor has direct application in scenarios were fast or even real-time acquisition is needed.

In this work we propose novel coding strategies that encode point location with subpixel accuracy. To this end we split the encoding and decoding phases into two steps: the first, called *phase coding/decoding*, maps the projector coordinate into a vector of values each with limited range, and the second, called *pattern generation* encodes the vectorial phase values onto physical quantities of
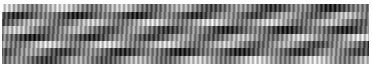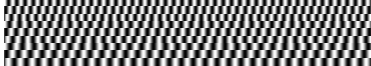
	Number-theoretic	Algebraic
Compound		
Subpattern		

Figure 1: Example patterns produced by the four proposed schemes.

the projected patterns. Phase coding strategies map the projector coordinate u into several wrapped phase values $\phi_i(u) \in [0; 2\pi)$ quantized into λ_i angular values. The coding strategy has to be invertible so that the projector coordinate can be recovered uniquely from the phase vector $\vec{\phi}(u) = (\phi_1(u), \ldots, \phi_n(u))$; further, in order to extract correspondences with subpixel precision, the map must be continuous, i.e., it should be able to encode and uniquely decode non-integral projector coordinates. On the other hand, the pattern generation step encodes the phase vector $\vec{\phi}(u)$ as observable physical quantities of the time series $f_u(t)$ of the projected light intensity.

The contribution of this paper is twofold. First, we propose and analyze two novel pattern generation methods (compound and subpattern) that allow to significantly reduce the number of projected patterns. Both schemes encode the phase vector as phases in a time-varying sinusoidal pattern: One simply encodes each coordinate in a separate subsampled sine wave, while the other encodes the whole vector as phases of the Fourier transform of the time-varying pattern. The second contribution is the introduction of a novel encoding technique which adopts a fine-to-coarse strategy allowing for smaller quantization at the fine level to increase precision, and larger quantization at the coarse level to reduce the total number of phases to be encoded. With the addition of a coding strategy derived from [3], we analyze a total of four different encoding schemes. Figure 1 shows examples of the patterns generated with the four schemes. The two rows show the pattern generation methods, while the columns show the coding strategies, with Number-theoretic referring to the scheme derived from [3], and Algebraic the newly proposed scheme.

The number-theoretic scheme presented in [3] uses period co-primality to map the phase vector onto a larger linear range. In particular, it projects sinusoidal patterns whose wavelengths are mutually co-prime, thereby allowing to deduce the overall position from the detected remainders. For example, we can encode 990 projector columns with three sinusoidal patterns of periods 9, 10, and 11 respectively. This way, column 382 will be encoded as phase $382 \bmod 9 = 4$ on the first sinusoid, phase $382 \bmod 10 = 2$ on the second, and phase $382 \bmod 11 = 8$ on the last sinusoidal wave. The algebraic approach we are proposing, on the other hand, adopts a hierarchical fine-to-coarse encoding which can selectively assign a larger number of samples to the low order representation to improve encoding precision. In this scheme we project different patterns each encoding one digit of the overall position. Thus, following the previous example, we can encode 1000 projector columns with three sinusoidal patterns each of period 10, and column 382 will be encoded as phase $382 \bmod 10 = 2$ on the first sinusoid, phase $382/10 \bmod 10 = 8.2$ on the second, and phase $382/100 \bmod 10 = 3.82$ on the last sinusoidal wave. The decimal part of each digit in the phase vector is maintained to ensure higher precision in the reconstruction. By changing the range of the phase vectors and the parameters of the pattern generation scheme, we can devise schemes with controllable number of patterns, either reducing them for fast reconstruction, or increasing them to either increase precision in low noise conditions or improve robustness to noise when the signal to noise ratio is low.

In conjunction the two contributions result in a coding strategy that can be tuned to be either more precise or more robust than the state-of-the-art for any fixed (small) number of projected patterns.

The rest of the paper is organized as follows: the next Section presents a review of the literature of range measurement by projected light; Section 3 introduces and analyzes the pattern generation strategies requiring a controllable number of projected patterns; following, Section 4 introduces a novel encoding scheme and Section 5 presents some experimental evaluation of the proposed coding schemes. Finally, Section 6 draws some conclusions.

2. Related Work

Among the most significant light-based approaches to 3D sensing, interferometry has certainly played a major role. Interferometric techniques for surface profiling are based on the superposition of multiple coherent electromagnetic wave-fronts resulting in an interferogram (or fringe pattern). The intensity distribution of the pattern varies as a function of the diffuse surface on which it is projected, and therefore can be utilized to undertake surface measurements. The relative height of the surface with respect to a reference plane can be retrieved from the analysis of an image of the modulated fringe by means of processing algorithms, among which Fourier transform profilometry (FTP) [4] has been by far the most widely studied.

Other widely adopted reconstruction techniques include structured light and phase-shifting methods. The simplest of these techniques is binary coding [5]; the method assigns to every pixel a codeword retrieved from the digitized sequence over time of projected black and white stripes, requiring $\log_2(n)$ pattern images to generate n code strings. Robustness of binary codes is improved by using Gray codes, where adjacent codes differ only in one bit. Both the techniques generate unique codes along each scanline, but at the same time are limited by their low resolution due to the inherently discrete nature of the coding. Also, the large number of projected patterns does not lead to an increased accuracy. Other structured light approaches make use of color-coded and grid-like patterns [5, 6].

Phase-shifting methods are based on the projection of periodic patterns with a given spatial period. Each projected pattern is obtained by spatially shifting the preceding one of a fraction of the period, and then captured by one or more cameras. The images are elaborated and the phase information at each pixel determined by means of M-step relationships [7]. Since the phase is distributed continuously within its period, phase-shifting techniques provide subpixel accuracy and achieve high measurement spatial resolution. Furthermore, the intensity change at each pixel for subsequent patterns is relative to

5

the underlying color and reflectance, which makes phase shift locally insensitive to texture variance to a certain degree. Again, in its basic formulation, phase-shifted structured light renders only relative phase values and thus it is ambiguous. When both an extended measuring range and a high resolution are required, a combined approach proves to be very powerful: the integration of Gray code and phase shift [8, 9] is an effective technique bringing together the advantages of both, providing disambiguation and high resolution, but the number of patterns to be projected increases considerably, and each strategy introduces a source of error [3].

A prevailing drawback within the phase shift class is related to an arbitrary shifting error [10] caused by the analog nature of the projecting device, which can influence the final evaluation of the phase values if proper compensation is not carried out. Nevertheless, the introduction of highly flexible digital video projection (DVP) technologies (namely LCD and DLP), and also by virtue of their availability and relatively low cost, lead to the complete avoidance of this kind of errors. However, this also resulted in geometrically less precise signals, and in addition for better visual effect usually both the light source and the imaging device generalize as nonlinear devices in such a way that the captured patterns deviate from the ideal sinusoid. This makes it problematic for the phase extraction step to yield accurate values, which ultimately leads to erroneous depth measurements [11, 12]. In [13] Pan *et al.* perform a theoretical analysis of the phase error induced by the non-sinusoidal waveforms and devise an iterative compensation algorithm taking into account the non-linearity due to the whole system, rather than the DVP effect alone. In general though, accuracy depends on the projector characteristics as a whole, as discussed in [14]. It has been demonstrated [7] that these errors can be somewhat alleviated by using a larger number of fringe patterns, however for dynamic applications requiring fast reconstruction (such as FTP and 3-step or 4-step profilometry [15, 16]) the non-sinusoidal phase error is considered as the dominant error source. As a matter of fact, non-linearity of the video projector is mainly due to gamma correction for image enhancement which inevitably decreases the accuracy and

6

resolution of the measurement. This problem has also been extensively studied [17, 18]. A common and simple schema is to gamma encode the patterns before projection; other approaches include iterative gamma-correction techniques based on statistical analysis of the captured fringes [19].

Recently, attention has shifted back to number-theoretical methods, whereas simpler strategies have been proposed. Lilienblum and Michaelis presented a simple and robust multi-period phase shift algorithm, which allows for very accurate and stable reconstructions [3]. In [20] a modification is given to the original Gushov-Solodkin algorithm; in their work, the relative phase values are checked and corrected before the unwrapping procedure takes place. The method is claimed to be faster and not influenced by inaccuracies in relative phase computation, however more importantly it is suited toward a simple implementation with off-the-shelf components and its parameters need fine-tuning in some measure.

The big advantages of the multi-period method presented in [3] is its relative simplicity and high efficiency. The phase-coded images can be directly employed in general stereo reconstruction systems, ensuring high quality and density of the code. Specifically, the lack of surface points in the final model is mainly due to occlusions and camera disparity, and measurement errors are very low thanks to the averaging and validation procedures implicit to the approach, that exclude a large percentage of errors and outliers before the actual surface reconstruction takes place. The main drawback lies in the fact that, typically, three or more phase-shift sequences are needed to entirely cover the projector range (typical values are 800 or 1024 projector pixels). This requires the projection of as many as three times more patterns than required with classical phase-shifting. Further, at least in the original formulation of Lilienblum and Michaelis, the phase increment in time and in space are correlated, requiring a larger number of patterns as the resolution of the projector increases.

In order to address the latter issue, and reduce the number of pattern required, in [21] a new compound phase-coding approach was presented. The main idea behind the Compound Phase Coding strategy was to project several

fringe patterns in a single spatio-temporal pattern, by encoding the phases of the fringe vector as phases of a Fourier term at different frequencies. This effectively decouples the spatial phase increments with the temporal increments used to create the patterns. However, the reduction in the dynamic range available to each frequency slot increases the noise of the estimated phase, and the Fourier composition forces the number of patterns projected to be an integral multiple of $2 * (k + 1)$ where k is the number of phase fringes used for coding.

3. Phase Coding and Pattern Generation

When adopting multiple phase coding strategies, the process of going from projector coordinates to the projected patterns can be divided into two main steps. The first step is that of *phase coding* and the dual step of *phase decoding*. Phase coding strategies map the projector coordinate u into several wrapped phase values $\phi_i(u) \in [0; 2\pi)$ quantized into λ_i angular values. The coding strategy has to be invertible so that the projector coordinate can be recovered uniquely from the phase vector $\vec{\phi}(u) = (\phi_1(u), \ldots, \phi_n(u))$; further, in order to extract correspondences with subpixel precision, the map must be continuous, i.e., it should be able to encode and uniquely decode non-integral projector coordinates. The second step is that of *pattern generation* and its dual *phase estimation*. The pattern generation step encodes the phase vector $\vec{\phi}(u)$ as observable physical quantities of the time series $f_u(t)$.

For example, the phase coding of the appraoch in [3] maps the projector coordinates onto a vector of remainders modulo mutually co-prime numbers, while the pattern generation simply maps each element on the vector onto the phase of a fixed-frequency and fixed-amplitude sinusoidal signal. The separation of the encoding process into phase coding and pattern generation is rather general and can be applied to approaches that are not limited to multiple phase shift. Gray codes in [8] fit perfectly the schema as they can be interpreted as dyadic angular quantizations (i.e., $\lambda_i = 2$), where phases are encoded using standard Gray codes and patterns are generated using the amplitude of Haar basis functions as the observable quantity rather than the phase of Fourier basis.

In this paper we propose two novel pattern generation/phase estimation methods that allow to significantly reduce the number of projected patterns. The first approach is a direct generalization of multi-period [3], which decouples spatial phase quantization with the temporal periods of the projected patterns, while the second approach encodes the m coordinates of the phase vector as the phases of the first m harmonics of the Fourier-transformed signal. In both cases, the decoupling of spatial and temporal periods allows us to choose where to invest in a larger number of projected patterns or signaling band to reduce the phase estimation error. Second, we introduce a novel algebraic encoding/decoding technique which adopts a fine-to-coarse strategy allowing for smaller quantization at the fine level to increase precision, and larger quantization at the coarse level to reduce the total number of phases to be encoded (and thus patterns to be projected). Compounded with the ability of the proposed pattern generation strategies of allocating extra patterns to specific phases, this results in a coding strategy that is both more precise and more robust than the state-of-the-art for any fixed (small) number of projected patterns. The combination of the two presented pattern generation strategies with the proposed algebraic encoding and the number-theoretical encoding of Lilienblum and Michaelis gives us four structured light coding strategies which will be analyzed for performance and robustness.

3.1. Phase Coding

In [3] the authors suggest to map projector coordinate u to the phase vector

$$\vec{\phi}(u) = (\phi_1(u), \ldots, \phi_n(u)) \quad \text{with} \quad \phi_i(u) = \frac{2\pi}{\lambda_i}(u \bmod \lambda_i), \qquad (1)$$

where $u \bmod \lambda_i$ is the remainder of the division of u by λ_i. In addition, the fringes of a pattern are assigned sequential natural numbers $\eta_i(u) \in \mathbb{N}$, which represent a simple counting of the fringes from left to right. A projector coordinate can then be directly obtained, for all $i = 1, 2, \ldots, n$, from a fringe number and a phase value:

$$u = (\eta_i(u) + \phi_i(u))\lambda_i. \qquad (2)$$

Since the only available values during decoding are λ and ϕ, the system of equations becomes ambiguous as the same value of u can be obtained for different values of η_i. This happens when two different projector coordinates $\tilde{u} \neq \tilde{\tilde{u}}$ yield the same phase values for all i. Under such conditions, the following derivation holds [3]:

$$\tilde{u} - \tilde{\tilde{u}} = (\eta_i(\tilde{u}) - \eta_i(\tilde{\tilde{u}}))\lambda_i, \tag{3}$$

and then, for some positive integers $x_1, x_2, ..., x_n$:

$$|\Delta u| = x_1\lambda_1 = x_2\lambda_2 = ... = x_n\lambda_n. \tag{4}$$

Therefore, a general condition is identified for generating unambiguous pattern sequences, by defining a maximum projector coordinate u_{max} up to which ambiguity can be excluded. Such a coordinate is the least common multiple of relatively prime periods λ_i, which for practical advantage shall entirely cover the projector range.

An efficient method is then given to calculate the fringe numbers from the ambiguous phase values at each pixel, taking advantage of a simple relationship between them. Given any pair of pattern sequences, for each image pixel it holds:

$$\lambda_i\phi_i(u) - \lambda_j\phi_j(u) = \lambda_j\eta_j(u) - \lambda_i\eta_i(u). \tag{5}$$

This makes it possible to construct a theoretical phase difference vector beforehand, and then use it to retrieve the fringe numbers when real phase measurements become available. In addition to providing an efficient way to obtain the fringe numbers, this method allows to assign each point a reliability value related to the deviation between measured and expected values. The use of theoretical phase difference vectors makes for a powerful test, which allows to identify erroneous or weak measurements (such as mixed phase values) caused, for instance, by sharp edges, involuntary object movements and light reflections. Once the unknown fringe numbers are calculated, projector coordinates can be easily retrieved for each pattern sequence with equation 2. Further, the independent measurements can be averaged to obtain the unwrapped phase at every pixel in an efficient way, leading to an increase in measurement accuracy.

3.2. Pattern Generation

Phase-shift based strategies use the phase of sinusoidal waves as quantities that can be recovered through the imaging process, but other approaches are possible; for example Gray codes use the amplitude of the signal as information carrier.

According to multi-period phase shift, the m coordinates of the phase vector $\vec{\phi}(u)$ are projected as m separate subseries $f_u^i(t)$, each of length λ_i, where each time series is a sinusoid of period λ_i:

$$f_u^i(j) = \cos(2\pi j/\lambda_i + \phi_i(u)) \,. \tag{6}$$

The final pattern f_u is then the juxtaposition of the patterns f_u^i, scaled and shifted to the projector intensity range (usually $[0; 255]$). Note that the shift only affects the level of the DC component and the amplitude of the higher frequency carriers, leaving the phase values unchanged.

3.3. Time-Decoupled Subpatterns

In the original formulation presented in [3] the periods λ_i represent both the phase increments as we move spatially through the projector coordinates and the time increments as we move along the pattern. This enforces a very strong requirement on the number of projected fringes which have to be $\sum_i \lambda_i$, where the λ_i must be co-prime and large enough that their product covers the whole projector range. Note, however, that the space and time increments are not really related and one can sample a full period of the sinusoidal pattern using any number of time-steps greater than 2. In general, we can encode the phase value $\phi_i(u)$ with k_i patterns as

$$f_u^i(j) = \cos\left(2\pi \frac{j}{k_i} + \phi_i(u)\right) \quad \text{with } i = 0, \ldots, k_i - 1 \,. \tag{7}$$

The sine and cosine of phase can then be simply estimated from the observed

11

temporal series $\bar{f}_u^i$ through Fourier analysis:

$$\rho \cos\left(\hat{\phi}_i(u)\right) \quad = \quad < \bar{f}_u^i | \cos > = \frac{1}{k_i} \sum_{j=0}^{k_i-1} \bar{f}_u^i(j) \cos\left(-2\pi\frac{j}{k_i}\right) \tag{8}$$

$$\rho \sin\left(\hat{\phi}_i(u)\right) \quad = \quad < \bar{f}_u^i | \sin > = \frac{1}{k_i} \sum_{j=0}^{k_i-1} \bar{f}_u^i(j) \sin\left(-2\pi\frac{j}{k_i}\right) , \tag{9}$$

from which we get the estimate

$$\hat{\phi}_i(u) = \arg\left(\rho\cos\left(\hat{\phi}_i(u)\right), \rho\sin\left(\hat{\phi}_i(u)\right)\right), \tag{10}$$

where $\arg(x,y)$ is the argument of complex number $x + iy$ and is in the range $[0; 2\pi)$.

To study the angular accuracy in the phase estimation, assume that the observed temporal series was subjected to a constant shift and rescaling and independent observation errors $\epsilon_0, \ldots, \epsilon_{k-1}$ of zero mean and constant standard deviation σ, i.e., $\bar{f}_u^i(j) = c + af_u^i(j) + \epsilon_j$. From this we have

$$E\left[\rho\right] \quad = \quad \frac{a}{2} \tag{11}$$

$$\mathrm{Var}\left(\rho\hat{\phi}_i(u)\right) \quad \approx \quad \mathrm{Var}\left(\rho\cos\left(\hat{\phi}_i(u)\right)\right) = \mathrm{Var}\left(\rho\sin\left(\hat{\phi}_i(u)\right)\right) =$$

$$= \quad \frac{\sigma^2}{k_i^2} \sum_{j=0}^{k_i-1} \cos^2\left(-2\pi\frac{j}{k_i}\right) = \frac{\sigma^2}{2k_i} \tag{12}$$

$$\mathrm{Var}(\hat{\phi}_i(u)) \quad \approx \quad \frac{\mathrm{Var}\left(\rho\hat{\phi}_i(u)\right)}{\mathrm{Var}(\rho)} = \frac{2\sigma^2}{k_i a^2} . \tag{13}$$

Thus, the angular error $\sigma_{\phi_i} = \frac{\sigma}{a}\sqrt{\frac{2}{k_i}}$ is inversely proportional to the signal-to-noise ratio a/σ and to the square root of the number of samples k_i. In this respect, using a number of samples proportional to the period cannot even be justified by the requirement that the phase error be the same for all subpatterns, as that would happen for a number of samples proportional to the *square* of the periods.
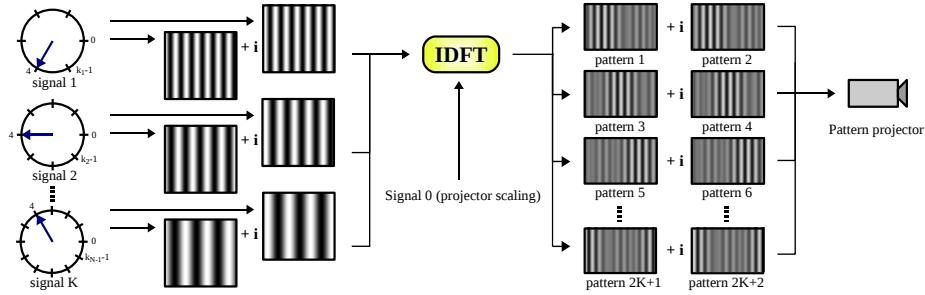
Figure 2: The composition of k fringe patterns, plus one unknown shift signal will produce a total of $2(k + 1)$ image patterns that will be projected onto the surface to be reconstructed. The shift pattern (projector scaling) accounts for the unknown value of the albedo of the surface.
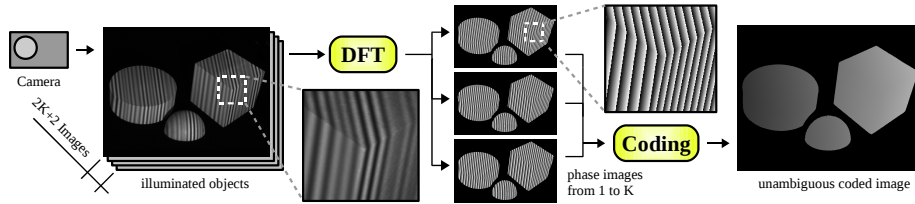

Figure 3: A total of 2k+2 images of illuminated objects are captured and single phase values are calculated for each composed fringe signal. Those values are subsequently used to get an unambiguous coding. Note that the intensity profile of each projected pattern is not sinusoidal.

3.4. Compound Pattern Generation

The compound phase coding approach was introduced in [21][1]. It encodes the m coordinates of the phase vector as the phases of the first m frequencies of the Fourier transform of the signal. Given a phase code $\vec{\phi} \in [0, 1)^k$, we create a $(k + 1)$-dimensional complex vector $\vec{x} \in \mathbb{C}^{k+1}$, where

$$
x_j = \begin{cases} 0, & \text{if } j = 0, \\ eq_j e^{-2\pi i \phi_j}, & \text{if } 1 \le j \le k. \end{cases} \tag{14}
$$

Here, $i = \sqrt{-1}$ and $\vec{eq} = (eq_1, \ldots, eq_k)$ with $\sum_{j=1}^{k} eq_j = 1$ is an equalization vector that indicates the relative strength of each multiplexed phase signal. Note that given x_j we can compute the phase as $\phi_j = \arg(\Re(x_j), \Im(x_j))$, where $\Re(z)$, $\Im(z)$ are the real and imaginary parts of the complex number $z \in \mathbb{C}$, respectively.

---

[1]Note that equations (14), (15), and (16) as well as Figures 2 and 3 are taken from [21].

Each complex number x_j represents the amplitude and phase of a sinusoidal component with frequency $\frac{j}{k+1}$ cycles per sample. Hence we can reconstruct the intensity sequence of that coordinate by computing the Inverse Discrete Fourier Transform of $\vec{x}$, obtaining the temporal pattern

$$f(n) = \frac{1}{k+1} \sum_{j=0}^{k} x_j e^{2\pi i \frac{j}{k+1} n}, \quad n = 0, \ldots, k.$$ (15)

We can then project separately the real and imaginary part of this vector as two time sequences obtaining a single set of $2(k+1)$ patterns to be projected to uniquely encode the phase vector $\vec{\phi}$ (see Figure 2).

The acquisition process introduces an additional linear deformation on the time series f_u, which depends on the physical properties of the object being scanned and which does not affect the phases. Let $\hat{f}_u(t)$ be the acquired gray-scale values and let $\vec{y} \in \mathbb{C}^{k+1}$ be its representation into a complex vector. The phase code $\vec{\phi}$ is finally recovered from $\vec{y}$ by computing the Discrete Fourier Transform, namely

$$\overline{x}_j = \sum_{n=0}^{k+1} y_n e^{-2\pi i \frac{n}{k+1} j},$$ (16)

and by extracting the argument from the resulting vector $\vec{\overline{x}} \in \mathbb{C}^{k+1}$ (see Figure 3).

This process allows to recover the phase code for each projector coordinate by taking only $2(k+1)$ measurements, where k is the number of signal periods. Nevertheless, one can also force a larger number of samples in order to increase accuracy, by appending null components to $\vec{x}$. More precisely, with M null components we need $2(M+k+1)$ measurements in order to recall the phase code $\vec{\phi}$.

It should be noted that a drawback of this approach is that encoding multiple signals in a single pattern reduces the effective projector intensity range available to encode each phase, increasing the effects of the discretization error and observation noise. Applying the same error analysis as the one performed for the time-decoupled subpatterns generation strategy, we see that the angular

14

error in the estimation of phase ϕ_i is

$$\sigma_{\phi_i} = \frac{\sigma}{a\,eq_i}\sqrt{\frac{2}{k}}, \tag{17}$$

where a is the total signal, eq_i is the equalization level assigned to phase ϕ_i and k is the total number of patterns projected. Hence, an m-wise increase in patterns assigned to each phase cannot balance the m-wise reduction in dynamic range available to each pattern. On the other hand, this coding strategy can encode m phases with as few as $2m + 1$ patterns, versus the minimum of $3m$ of the subpattern strategy.

4. Algebraic Coding

In the phase encoding strategy we adopt a fine-to-coarse approach where the first phases have short spatial periods, while subsequent phases have larger periods; the last one has a spatial period larger than the maximum projector coordinate. Let $\lambda_1, \ldots, \lambda_m$ be phase quantizations with the only requirement that $\prod_{i=1}^{m} \lambda_i \geq M$, with $M \leq u_{max}$ the maximum projector coordinate. We encode the projector coordinate u into the phase vector $\vec{\phi}(u) = (\phi_1(u), \ldots, \phi_m(u))$ where

$$\phi_i(u) = 2\pi \operatorname{frac}\left(\frac{u}{\prod_{j=1}^{i} \lambda_i}\right) \tag{18}$$

where $\operatorname{frac}(x) = x - \lfloor x \rfloor$ is the fractional part of $x \geq 0$. The name *algebraic coding* derives from the fact that the encoding is constructed in a similar manner to the way integers are represented using the standard positional notation. Assume that phase ϕ_i is divided into λ_i angular bins, and assume that the projector coordinate u is represented in positional notation as the sequence of digits $d_1, \ldots, d_m$, where position i uses base λ_i, i.e., d_i is in the range $0, \ldots, \lambda_i - 1$. Then the digit d_i indicates the angular bin in which ϕ_i falls, and the integral coordinate u can be recovered by $u = \sum_{i=1}^{m} d_i \prod_{j=1}^{i} \lambda_i$. Thus, using this algebraic representation, we can reconstruct the integral coordinate by looking at the angular bucket in which the phase estimation falls. In practice, the projector coordinate is estimated with subpixel precision, but the fractional part of

15

the location is extracted only from the first phase, which is localized with the highest precision.

In particular, the decoding phase is as follows: first we get the low-order digit with subpixel precision $d_1 = \hat{\phi}_1 \frac{\lambda_i}{2\pi}$ and set the current estimate at $h_1 = d_1$. Then the estimate is refined adding the other phases according to the following recurrence:

$$d_{i+1} \quad = \quad \text{round}\left(\hat{\phi}_{i+1} \frac{\lambda_{i+1}}{2\pi} - \frac{h_i}{\prod_{j=1}^{i} \lambda_i}\right) \tag{19}$$

$$h_{i+1} \quad = \quad d_{i+1} \prod_{j=1}^{i} \lambda_i + h_i. \tag{20}$$

The final estimate h_m is the estimated projector coordinate.

This coding strategy allows us to completely decouple the quantization levels λ_i, which can be any value as long as their product is larger than the maximum coordinate. Further, since the information about subpixel localization is extracted only from the low-order bit, we can invest in it both by reducing the number of quantization bins, and by increasing the number of samples or the equalization level assigned to the first phase. This separates it from other hierarchical methods such as Gray coding since this equalization can reduce the effect of noise on the final subpixel precision. In addition, the continuity of high-order bins eliminates the noise due to the spatial quantization effect. Figure 1 shows an example set of patterns obtained using the proposed pattern generation strategies in conjunction with the number-theoretic and algebraic coding schemes.

4.1. Phase Correction

Phase shift methods achieve subpixel accuracy assuming that by integrating the contribution of adjacent projector pixels one obtains a phase that is a linear combination of the phases projected at the two projector locations. Specifically, the assumption is that if a camera pixel integrates projector pixel u for a proportion α and projector pixel $u + 1$ for a proportion $1 - \alpha$, then the observed phase is $\hat{\phi} \approx \alpha\phi(u) + (1 - \alpha)\phi(u + 1)$. However, the reconstruction obtained

16

$e^{i\left(\phi_{\text{int}}+\frac{2\pi}{\lambda}\right)}$

$e^{i(\phi_{\text{int}}+\phi_f)}$

$1-\alpha$

α

$e^{i\phi_{\text{int}}}$
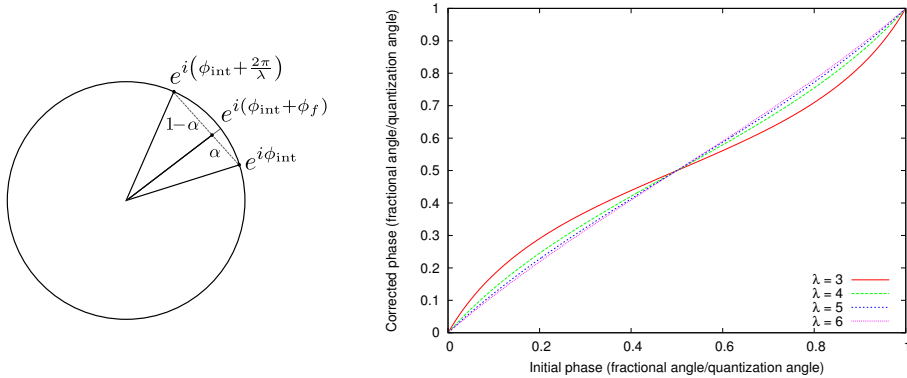
(a) Linear interpolation of the phase/amplitude vectors. The interpolated angle ϕ_f is different from the ideal linearly interpolated value $\alpha\frac{2\pi}{\lambda}$.

(b) Effect of the phase correction for angular quantization ranging from 3 to 6. Measures are in initial/corrected angles over minimum quantization angle $(2\pi/\lambda)$.

Figure 4: The linear interpolation of the phase/amplitude vectors at adjacent quantized angles does not result in a linear interpolation of the angles. This requires a correction for the fractional part of the phase.

from the superposition of the signals interpolates linearly the phase/amplitude vectors, which does not result in a linear interpolation of phase angles (see Figure 4.a). This results in an estimation error of the subpixel part of the code that is more severe as the phase quantization is coarser. This source of error is usually not a problem with the number-theoretical multi-phase shift method as, in practice, for common projectors the phases are quantized in 9 or more levels, and thus errors incurred with the linear approximation are limited. In our case, however, we aim at reducing the low-order quantization as much as possible in order to increase subpixel precision of the algebraic coding, resulting in severe linearization errors. We avoid this by correcting the estimated phase taking into account the non-linearity of the angular interpolation. Let

$$\phi_{\text{int}} = \lfloor\frac{\phi}{2\pi}\lambda\rfloor\frac{2\pi}{\lambda} \tag{21}$$

$$\phi_f = \text{frac}\left(\frac{\phi}{2\pi}\lambda\right)\frac{2\pi}{\lambda} \tag{22}$$

be the integral and fractional part of the quantization of phase ϕ into λ angular bins, and let a be the signal amplitude at this location. It is worth reminding

17

that one of the main assumptions is that the projected signal amplitude is constant, while the received signal amplitude depends on the surface orientation and albedo, but it is constant at any given location. The reconstructed phase amplitude vector is

$$(1 - \alpha)ae^{i\phi_{\text{int}}} + \alpha ae^{i\left(\phi_{\text{int}}+\frac{2\pi}{\lambda}\right)} = ae^{i\phi_{\text{int}}}\left((1 - \alpha) + \alpha e^{i\frac{2\pi}{\lambda}}\right) = ae^{i\phi_{\text{int}}}be^{i\phi_f} \ , \ (23)$$

where b is an amplitude correction term due to the amplitude reduction caused by the linear interpolation. Note that this reduction in the amplitude of the coding frequencies means that at low quantization level we are trading signal strength for angular discrimination, and this could be counterproductive on low signal conditions.

From equation (23) we obtain $be^{i\phi_f} = (1 - \alpha) + \alpha e^{i\frac{2\pi}{\lambda}}$, normalizing the phase/amplitude vector and looking at its real part, we get the following relation for the observed fractional part of the phase:

$$\cos(\phi_f) = \frac{(1 - \alpha) + \alpha \cos\left(\frac{2\pi}{\lambda}\right)}{\sqrt{\left((1 - \alpha) + \alpha \cos\left(\frac{2\pi}{\lambda}\right)\right)^2 + \alpha^2 \sin^2\left(\frac{2\pi}{\lambda}\right)}} \tag{24}$$

Solving for α and applying the recovered interpolation proportion to the quantization angles, we obtain the corrected phase:

$$\phi_{\text{corrected}} = \phi_{\text{int}} + \alpha \frac{2\pi}{\lambda} = \phi_{\text{int}} + \frac{\sin(\phi_f)}{\sin(\phi_f) - \sin\left(\phi_f - \frac{2\pi}{\lambda}\right)} \frac{2\pi}{\lambda} \tag{25}$$

Figure 4.b plots the effect, for small angular quantizations, of applying the phase correction to the fractional part of the phase.

5. Experimental Results

In order to validate the proposed techniques we performed both synthetic and real world experiments. In the first set we assessed the effect of noise, using different coding strategies and parameters, against a ground-truth value for the reconstructed projector coordinates. In the second set we compared the codes obtained with the proposed approach on a test rig for structured light techniques that has been internally developed in our lab. Here we did not have

ground-truth, so we compared our measurements with those given by the Multi-Period Phase Shift proposed in [3], where the time dimension was significantly up-sampled to reduce the effects of noise. In particular, a total of 90 patterns were projected, as compared to the minimum possible with this strategy of 9 patterns. The choice of projecting 90 patterns was driven by the assumption that by increasing the number of patterns by an order of magnitude, the errors in the reconstruction of the (subpixel) phase values and thus of the final column code will be similarly reduced to a level much smaller than the decoding error incurred with the schemes under analysis.

Number-theoretical encoding using periods $7, 10, 11$.

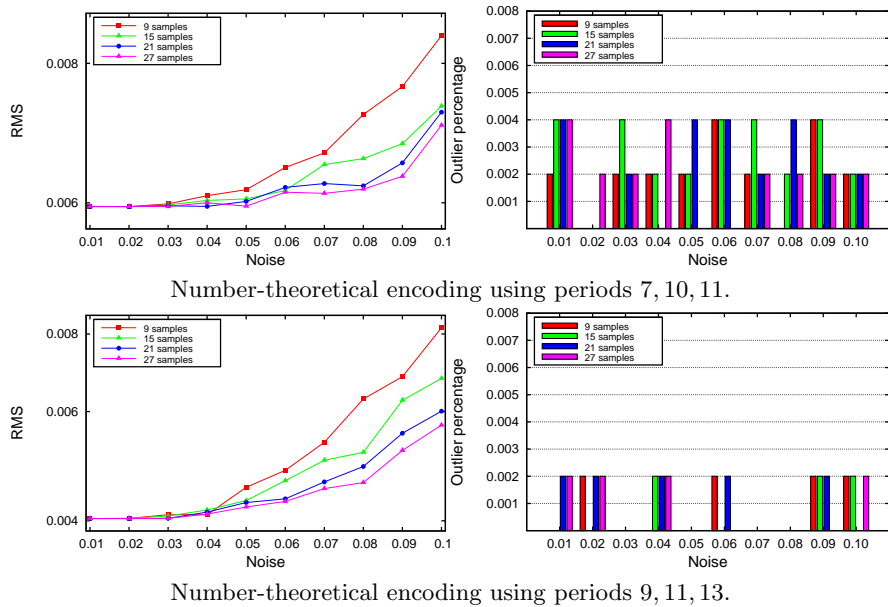Number-theoretical encoding using periods $9, 11, 13$.

Figure 5: Precision and robustness to noise of the number-theoretical encoding strategy in conjunction with the subpattern generation strategy. Left: code RMS as a function of the standard deviation of the noise. Right: percentage of outliers (codes with error grater than 1 projector pixel) as a function of the standard deviation of the noise. Same color represents same coding scheme.

5.1. Synthetic Experiments

To assess the robustness of the coding strategies with respect to noise, we performed a set of synthetic experiments, where an increasing amount of random additive noise was added to the patterns before decoding, and then the

19

error between the exact and the estimated coordinate was computed. The patterns for non-integral location u were generated by linearly mixing the patterns obtained at locations $\lfloor u \rfloor$ and $\lceil u \rceil$. For each approach we applied a selection of quantization values corresponding to various levels of the quality/speed trade-off. For the number-theoretical (NT) approaches we used periods $7, 10, 11$ and $9, 11, 13$ which are common values satisfying the constraint of relative primality, that offer similar quantization errors for all phases, and provide a sufficient number of distinct codes for a normal projector (770 codes the first, 1287 the second). For the algebraic coding we used periods $8, 10, 10$ and $5, 13, 13$. the first provides 800 codes with almost equal quantization error for low and high bits, while the second provides 845 codes with a much coarser quantization for the lowest bits, which should result is higher subpixel accuracy. For the approaches using the subpattern generation strategy, we used an increasing number of time samples evenly distributed among the phases, namely 9, 15, 21, and 27. Further, for the algebraic coding we tested the advantage of unevenly distributing the samples assigning more to the low-order phase. To this end we tested the sample distributions $(3, 3, 3)$, $(7, 4, 4)$, $(11, 5, 5)$, and $(15, 6, 6)$, ranging from the lowest possible number of patterns, to a very high number of patterns combined with a large up-sampling of the low-order phase. Clearly there is no guarantee for the optimality of these values but they do provide a set of reasonable values spanning the range of the quality/speed tradeoff. Table 1 shows the projected patterns corresponding to this selection of parameters.

For each test we drew 500 locations in the range $[0; 100]$ and computed the root mean square (RMS) error between the location and its estimation after the encoding/decoding process.

Figures 5 and 6 plot the results using the number-theoretical and the algebraic encoding respectively. The plots on the left show the RMS as a function of the standard deviation of the noise assuming a unit signal amplitude. The RMS was computed among locations that decoded within one projector pixel. Larger errors are caused by a phase estimation outside of the correct bin and should be considered as outliers. The plots on the right show the percentage of outliers
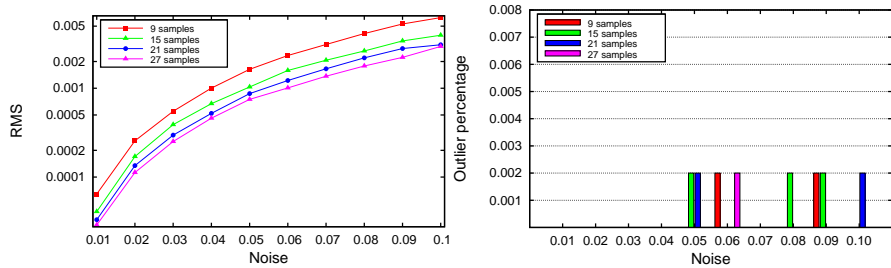
| Compound		
NT 7,10,13		
NT 9,11,13		
Alg 5,13,13		
Alg 8,10,10		
Subpattern		
NT 9 samples		
NT 27 samples		
Alg 3,3,3 samples		
Alg 15,6,6 samples		
Alg 21,3,3 samples		

Table 1: Actual sequences of patterns projected accordingly to the tested compound and subpattern coding strategies.
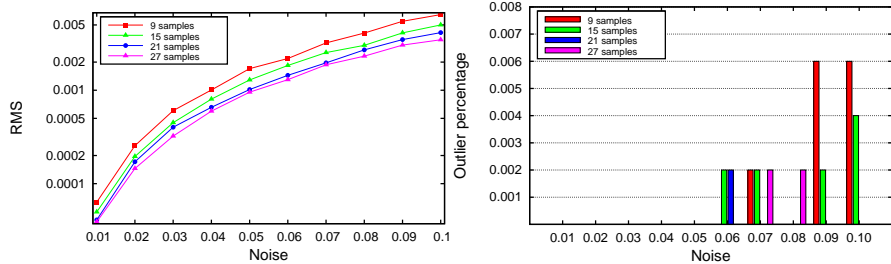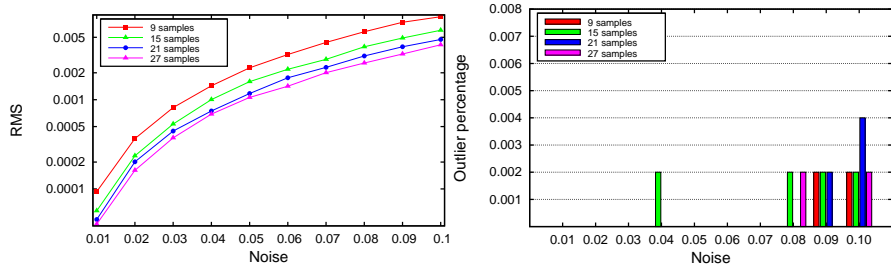
Algebraic encoding using quantizations $5, 13, 13$ and evenly distributed samples.

Algebraic encoding using quantizations $5, 13, 13$ and unevenly distributed samples.
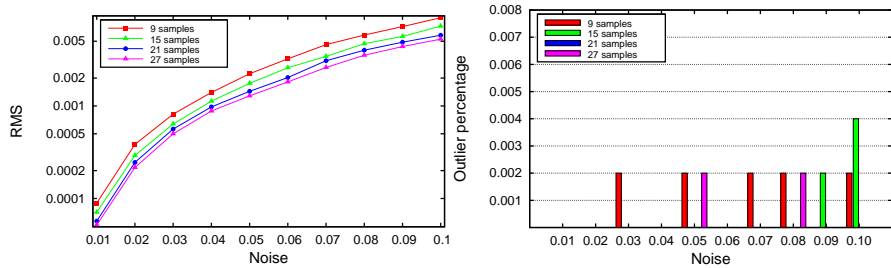
Algebraic encoding using quantizations $8, 10, 10$ and evenly distributed samples.

Algebraic encoding using quantizations $8, 10, 10$ and unevenly distributed samples.

Figure 6: Precision and robustness to noise of the algebraic encoding strategy in conjunction with the subpattern generation strategy. Left: code RMS as a function of the standard deviation of the noise. Right: percentage of outliers (codes with error grater than 1 projector pixel) as a function of the standard deviation of the noise. Same color represents same coding scheme.

Number-theoretical encoding using periods $7, 10, 11$.

Number-theoretical encoding using periods $9, 11, 13$.

Algebraic encoding using quantizations $5, 13, 13$.
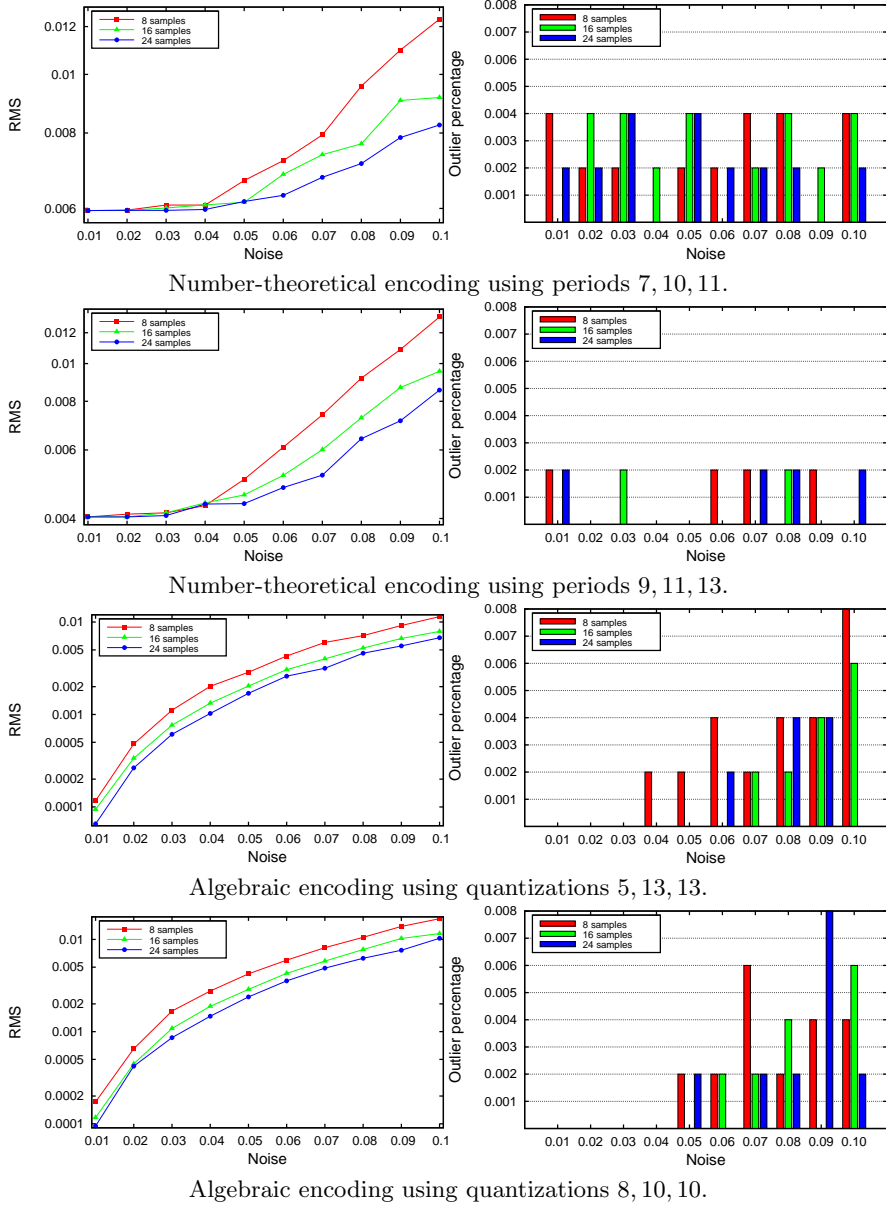

Algebraic encoding using quantizations $8, 10, 10$.

Figure 7: Precision and robustness to noise of the encoding strategies under study in conjunction with the compound pattern generation strategy. Left: code RMS as a function of the standard deviation of the noise. Right: percentage of outliers (codes with error grater than 1 projector pixel) as a function of the standard deviation of the noise. Same color represents same coding scheme.

and assess the robustness of the coding scheme. From the plots we can see that in all cases a larger number of samples imply higher precision and robustness. Further, we see that with the number-theoretical encoding the RMS grows at a faster rate than using the algebraic coding, while still exhibiting similar robustness. This suggests that the increase in angular discrimination provided by the low quantization of the low-order phase of the algebraic coding results in higher subpixel precision even in the presence of noise, without resulting in a higher number of outliers. As for the effect of varying the quantization level or redistributing the samples in algebraic coding, as expected a smaller quantization or a redistribution of samples in favor of the low-order phase reduce the RMS but also increase the number of outliers.

Figure 7 plots the RMS obtained using the compound pattern generation strategy. Here various sampling levels were obtained by up-sampling the time-series by a factor of 2 and 3, resulting in 8, 16 and 24 projected patterns. The algebraic coding confirms the better scaling with respect to noise, but it also results in a smaller rate of outliers, at least for small noise levels.

5.2. Real World Experiments

All the following experiments have been run on a test rig for structured light techniques that has been internally developed in our lab. The rig is made up of a motorized plate for object positioning, four cameras and an illumination source mounted on a motorized liftable platform. Specifically the cameras are equipped with a 1/2 inch CMOS sensor which offers a full 1280x1024 resolution. The cameras are monochrome, thus no Bayer filters are placed over the sensor. While four cameras are available, in this experiment set we use only one to test the codes. The illumination source is a 800x600 color DLP projector which we use to project the monochromatic patterns. The system is controlled by a standard PC housed into the base of the rig. This PC is a 2.8 GHz AMD quad core system with 2 Gigabytes of ram.

In Table 2 we show the performance of different compound coding strategies when dealing with a set of real-world objects that exhibit wide variations
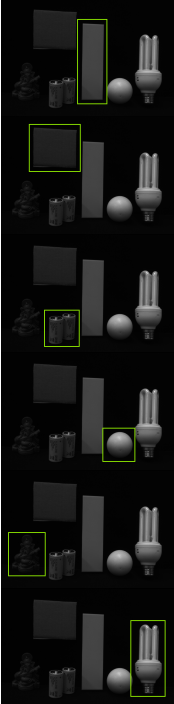
Compound						
	Strategy	Points	Error	max	Displacement	max
	NT 7,10,13	55571	0.302 ± 10.409	559.9	0.042 ± 0.027	0.662
	NT 9,11,13	55585	0.519 ± 17.734	935.8	0.042 ± 0.027	0.642
	Alg 5,13,13	55617	0.070 ± 0.985	65.1	0.039 ± 0.025	0.785
	Alg 8,10,10	55652	0.036 ± 0.065	8.156	0.038 ± 0.025	0.723
	NT 7,10,13	70984	4.327 ± 42.956	650.5	0.061 ± 0.039	3.981
	NT 9,11,13	71112	7.842 ± 73.689	1143.7	0.063 ± 0.037	1.256
	Alg 5,13,13	70993	0.207 ± 2.970	65.2	0.055 ± 0.032	0.504
	Alg 8,10,10	71465	0.079 ± 0.061	0.464	0.081 ± 0.045	0.534
	NT 7,10,13	17495	4.753 ± 47.132	650.6	0.054 ± 0.038	0.436
	NT 9,11,13	17504	5.546 ± 59.863	935.6	0.056 ± 0.040	0.496
	Alg 5,13,13	17553	0.217 ± 3.252	65.1	0.043 ± 0.029	0.327
	Alg 8,10,10	17667	0.066 ± 0.053	0.559	0.069 ± 0.046	0.556
	NT 7,10,13	11705	0.402 ± 11.209	350.4	0.038 ± 0.027	0.561
	NT 9,11,13	11721	0.873 ± 20.246	792.2	0.037 ± 0.028	0.752
	Alg 5,13,13	11814	0.054 ± 0.074	4.890	0.041 ± 0.029	0.477
	Alg 8,10,10	11792	0.046 ± 0.755	80.1	0.033 ± 0.025	0.439
	NT 7,10,13	8072	54.511 ± 152.075	819.5	0.126 ± 0.192	6.424
	NT 9,11,13	9107	71.647 ± 217.477	1143.8	0.130 ± 0.190	4.166
	Alg 5,13,13	8363	13.417 ± 91.544	785.1	0.094 ± 0.162	4.452
	Alg 8,10,10	12046	0.419 ± 10.076	720.1	0.084 ± 0.095	3.774
	NT 7,10,13	41652	0.922 ± 20.998	559.9	0.035 ± 0.024	0.261
	NT 9,11,13	41674	0.984 ± 21.214	494.8	0.035 ± 0.024	0.416
	Alg 5,13,13	41799	0.079 ± 0.212	5.206	0.051 ± 0.033	0.499
	Alg 8,10,10	41806	0.061 ± 0.048	0.529	0.050 ± 0.033	0.516

Table 2: Coding performance of the compound pattern generation strategy on objects with different albedo. The values are coding errors and thus are expressed in (dimensionless) coding units.

both in albedo and shape. Specifically, we used two algebraic and two number-theoretical pattern sequences. As in the synthetic experiments, we chose period sets of length 7,10,13 and 9,11,13 respectively for the number-theoretical tests and two quantizations of 5,13,13 and 8,10,10 for the algebraic coding. For each experimental condition we projected both horizontal and vertical pattern sequences, thus obtaining a bi-dimensional code for each image pixel. This code is used to compute two measures of deviation from the ground-truth. The first measure is the difference between the code obtained at a particular image pixel with respect to the ground-truth code obtained projecting 90 patterns; we call this quantity *code error*. The second measure used is the *code displacement*, i.e., the Euclidean distance between a point in the ground-truth image and the subpixel location containing the same code by bilinear interpolation in the
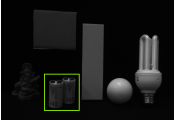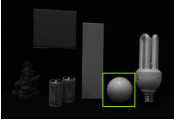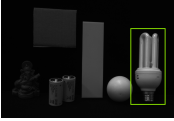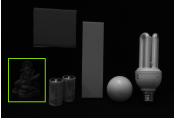
Subpattern						
	Strategy	Points	Error	max	Displacement	max
	NT 9 samples	56030	0.035 ± 0.027	0.351	0.033 ± 0.023	0.820
	NT 27 samples	56015	0.012 ± 0.010	0.189	0.016 ± 0.012	0.443
	Alg 3,3,3 samples	56317	0.040 ± 0.029	0.318	0.041 ± 0.025	0.544
	Alg 15,6,6 samples	56278	0.015 ± 0.011	0.132	0.017 ± 0.011	0.293
	Alg 21,3,3 samples	56313	0.013 ± 0.010	0.105	0.018 ± 0.011	0.316
	NT 9 samples	72515	0.128 ± 6.563	792.2	0.041 ± 0.024	0.354
	NT 27 samples	72504	0.019 ± 0.014	0.140	0.022 ± 0.013	0.193
	Alg 3,3,3 samples	72658	0.052 ± 0.040	0.396	0.044 ± 0.030	0.301
	Alg 15,6,6 samples	72646	0.026 ± 0.019	0.139	0.027 ± 0.015	0.183
	Alg 21,3,3 samples	72660	0.018 ± 0.014	0.127	0.023 ± 0.013	0.166
	NT 9 samples	17936	0.233 ± 9.810	494.8	0.050 ± 0.025	0.247
	NT 27 samples	17935	0.017 ± 0.014	0.118	0.023 ± 0.016	0.150
	Alg 3,3,3 samples	17996	0.049 ± 0.038	0.301	0.050 ± 0.032	0.277
	Alg 15,6,6 samples	17986	0.022 ± 0.016	0.148	0.023 ± 0.017	0.191
	Alg 21,3,3 samples	17992	0.015 ± 0.012	0.126	0.021 ± 0.015	0.166
	NT 9 samples	11982	0.075 ± 4.557	494.8	0.042 ± 0.021	0.451
	NT 27 samples	11981	0.011 ± 0.010	0.108	0.016 ± 0.013	0.235
	Alg 3,3,3 samples	12031	0.046 ± 0.033	0.353	0.042 ± 0.027	0.290
	Alg 15,6,6 samples	12027	0.028 ± 0.014	0.155	0.016 ± 0.014	0.221
	Alg 21,3,3 samples	12031	0.020 ± 0.014	0.133	0.016 ± 0.013	0.199
	NT 9 samples	41958	0.194 ± 7.925	494.7	0.026 ± 0.018	0.322
	NT 27 samples	41958	0.022 ± 0.010	0.116	0.012 ± 0.009	0.102
	Alg 3,3,3 samples	42073	0.040 ± 0.031	0.266	0.035 ± 0.022	0.271
	Alg 15,6,6 samples	42073	0.024 ± 0.016	0.158	0.015 ± 0.011	0.217
	Alg 21,3,3 samples	42077	0.025 ± 0.015	0.120	0.014 ± 0.011	0.168
	NT 9 samples	26300	6.339 ± 65.952	743.6	0.065 ± 0.039	0.419
	NT 27 samples	26248	0.029 ± 0.023	0.151	0.036 ± 0.021	0.155
	Alg 3,3,3 samples	27500	0.080 ± 0.062	0.403	0.082 ± 0.046	0.315
	Alg 15,6,6 samples	27516	0.036 ± 0.027	0.176	0.041 ± 0.024	0.207
	Alg 21,3,3 samples	27474	0.028 ± 0.022	0.163	0.038 ± 0.022	0.186

Table 3: Coding performance of the subpattern generation strategy on objects with different albedo. The values are coding errors and thus are expressed in (dimensionless) coding units.

experimental image. We can regard this quantity as the displacement error that would be committed when the coding is used in matching or stereo correspondence applications. In the columns of Table 2 we show the object used for the measurements, the number of successfully coded points, and finally the average, standard deviation, and maximum value of both the code error and displacement. As expected, compound coding works more reliably with smooth, uniform objects and Lambertian surfaces, such as the rectangular plastic strip, the white wooden sphere and, to some level, the white lamp. With those objects the performance obtained with the number-theoretical strategy is worse

than that obtained adopting the algebraic coding. This is due to the completely wrong coding retrieved from difficult boundary points, as can be noticed by looking at the large maximal errors. The coding error is even more dramatic with more complex or darker objects, such as the pair of batteries and the small statue. This is due to the decrease in dynamic range produced by the compound approach, and affects both the number-theoretical strategy and the proposed algebraic coding. However we can see that, despite using the same low number of patterns, algebraic coding exhibits far better performance and is able to cope with the most difficult objects exhibiting only a limited reduction in accuracy.

We repeated the same set of experiments with separately projected patterns. The sequences of patterns used and the results obtained are shown respectively in Table 1 and in Table 3. It is immediate to see that the overall error obtained is significantly lower than the one obtained with the compound strategies. The only gross coding errors happen with the number-theoretical approach when using only 9 patterns (three for each period). When using 27 samples the results are about one order of magnitude more accurate and we observe no mis-codings at all. It should be noted that the algebraic coding technique does not incur in any coding error even when using only 9 patterns. Furthermore, the ability to selectively redistribute the additional patterns to the low-order phase allows for a significant improvement in precision without incurring the risk of serious mis-codings.

In Figure 8 the multi-object scene used for the experimental validation is shown alongside some examples of the coding obtained with each strategy presented in this paper. The first two columns show the compound techniques. The number-theoretical one (first column) is able to recover correctly only a limited portion of the overall illuminated points and the algebraic technique allows to assign a valid code to a few more. However, by looking at the zoomed area, we can see that the former suffers more from gross coding outliers. By converse, the non-compound approaches offer a significantly higher number of coded points. In particular, the algebraic coding (fourth column) does not contain any outlier, while the number-theoretical strategy (third column) allows some to slip in. In
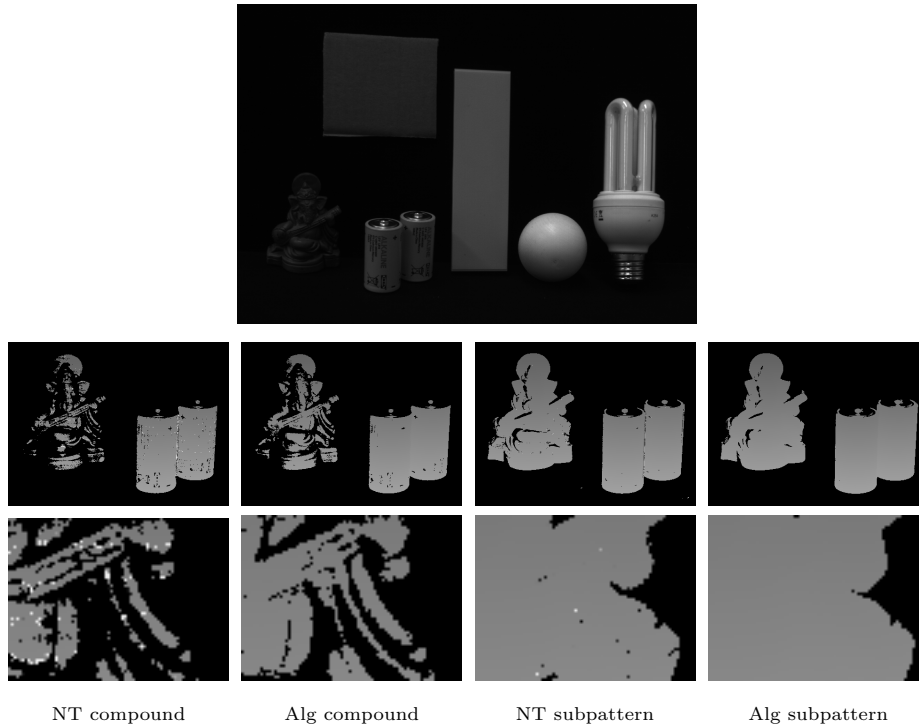
27

Figure 8: Test scene and example code images with respective close-ups for the number-theoretical and algebraic coding. Grey levels represent the detected code, from 0 (black) to 800 (white). Uncoded pixels are black.

the following we will explore these phenomena further.

In Figure 9 we show the effects of different coding parameterizations on the average and maximum code difference. In the first row we analyze the compound pattern generation techniques. In these experiments we test several levels of up-sampling ranging from no up-sampling (U1), to twice the number of patterns (U2) to three times the number of patterns (U3).

For each level we show the number-theoretical technique (NT), which obtains always the worst results regardless of the up-sampling level, and the algebraic coding scheme. The latter is tested with different equalization profiles: specifically level Eq1 assigns equal signal to all the digits, while Eq2 and Eq3 give respectively more signal to the high-order and low-order phases. Equalization does not affect the results much, although assigning more weight to the more
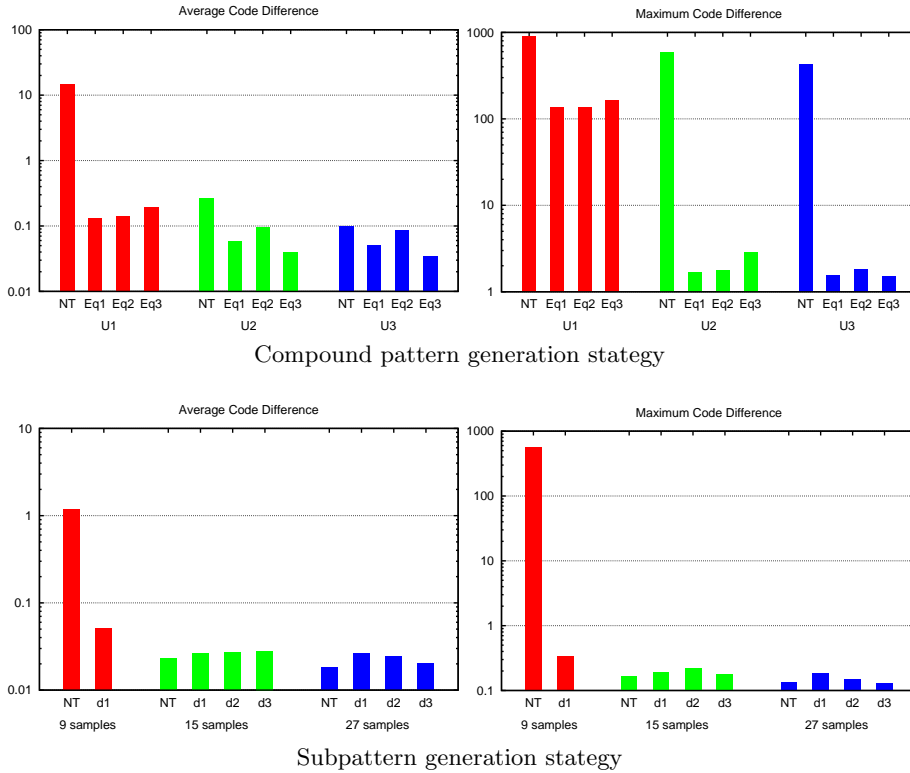
Figure 9: Accuracy of the coding with respect to different parameterizations of the coding schemes.

significant digits seems to result in a less accurate coding. In the second row we test the techniques based on separate subpatterns. Here we executed both the number-theoretical and the algebraic coding strategies with a number of patterns ranging from 9 to 27 samples. While it is apparent that the accuracy of the number-theoretical scheme improves with an increasing number of projected patterns, it should be noted that the algebraic coding scheme allows to obtain reasonable results even with just 9 patterns. When using more patterns we test three different assignments of the additional samples: with distribution d1 all the patterns are evenly divided for each phase period, while distributions d2 and d3 are progressively more biased toward the low-order phase. While the influence is not very strong, it is clear that giving more signal to the low-order phase leads to improved accuracy. Overall, these experiments show that the NT

technique does not works well with the compound pattern generation strategy, in addition, even with the subpattern strategy it perform poorly when the number of projected patterns is close to the Nyquist limit. Conversely it performs marginally better than the algebraic method in most situations when used with the subpattern strategy and a large number of patterns.

5.3. Surface Reconstruction

The previous set of experiments gives an extensive quantitative analysis of the robustness and accuracy of the coding strategies proposed. However, it is difficult to relate such numerical data to their actual influence on the quality of reconstructed surfaces. For this reason we also produced a set of 3D range images that provide qualitative evaluation of the influence of the coding scheme to the reconstruction process. Specifically, we used our test rig to obtain range images from two subjects that pose the most significant hurdles to the reconstruction process: The first is the light bulb, that exhibits reflections and partial transparency; the second is the Ganesha figurine, which has a very low albedo and presents several self-occlusions.

In Fig. 10 we show the reconstructions obtained from the light bulb. In the first row the surface obtained with the "Ground Truth" method is shown, i.e., the reference method also used for the quantitative experiments: the Number Theoretical technique with a very large number of projected patterns. For each method we also present two zoomed area (indicated with red boxes) that highlight a boundary zone and a flat area respectively. This is meant to evaluate the behavior of the techniques around the edges of the objects and their smoothness.

With the light bulb there is no significant difference in the number of triangulated points among the various methods. This is probably due to the high albedo of the subject and thus to the good level of signal received by the cameras. However, the Algebraic compound technique seems to loose a few points on the fluorescent tubes. In general, the subpattern methods allow to triangulate more points in the screw part of the bulb, which is mostly reflective. It is interesting to note that the Number Theoretical method used with the sub-
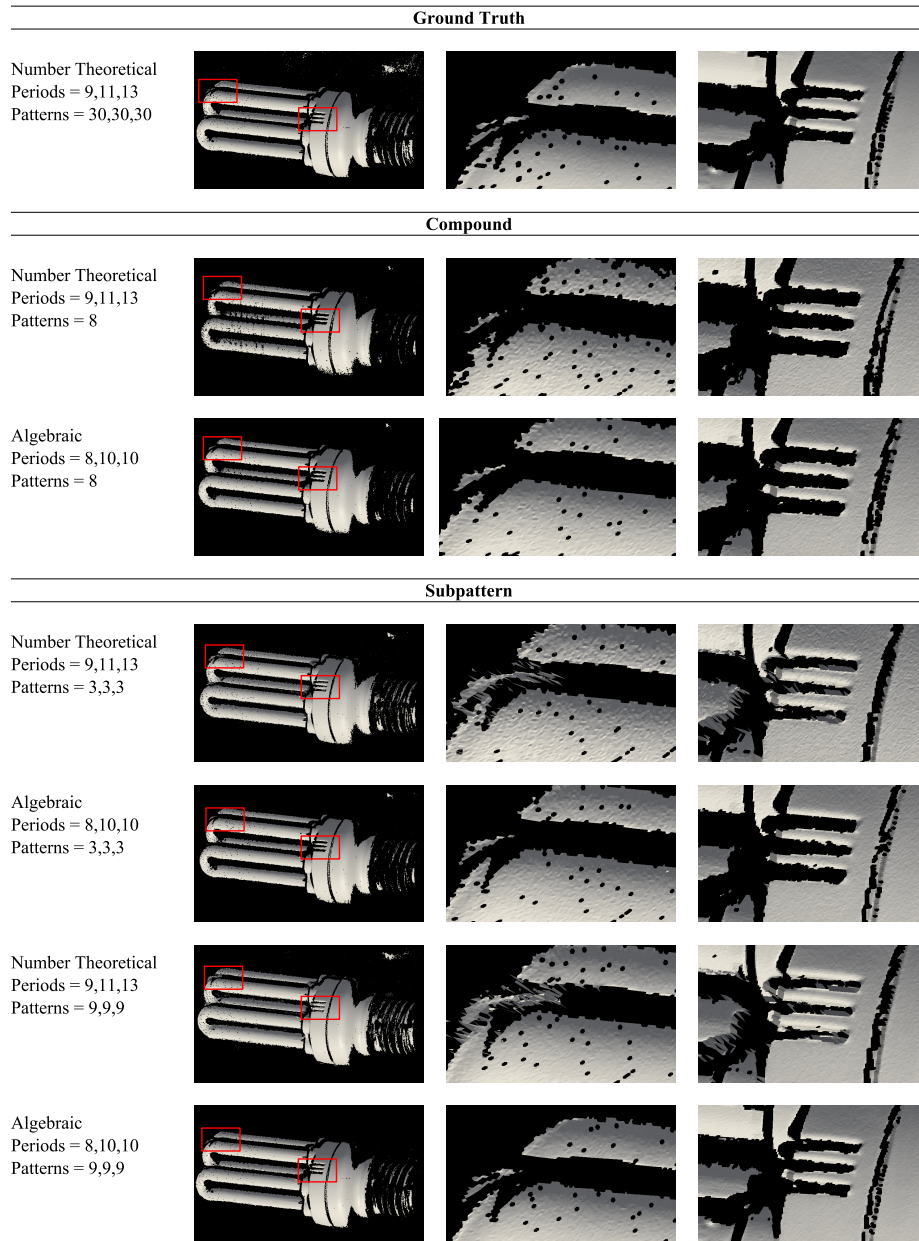
Figure 10: Qualitative analysis of the surfaces obtained from the light bulb subject.

Ground Truth

Number Theoretical
Periods = 9,11,13
Patterns = 30,30,30

Compound

Number Theoretical
Periods = 9,11,13
Patterns = 8

Algebraic
Periods = 8,10,10
Patterns = 8

Subpattern

Number Theoretical
Periods = 9,11,13
Patterns = 3,3,3

Algebraic
Periods = 8,10,10
Patterns = 3,3,3

Number Theoretical
Periods = 9,11,13
Patterns = 9,9,9

Algebraic
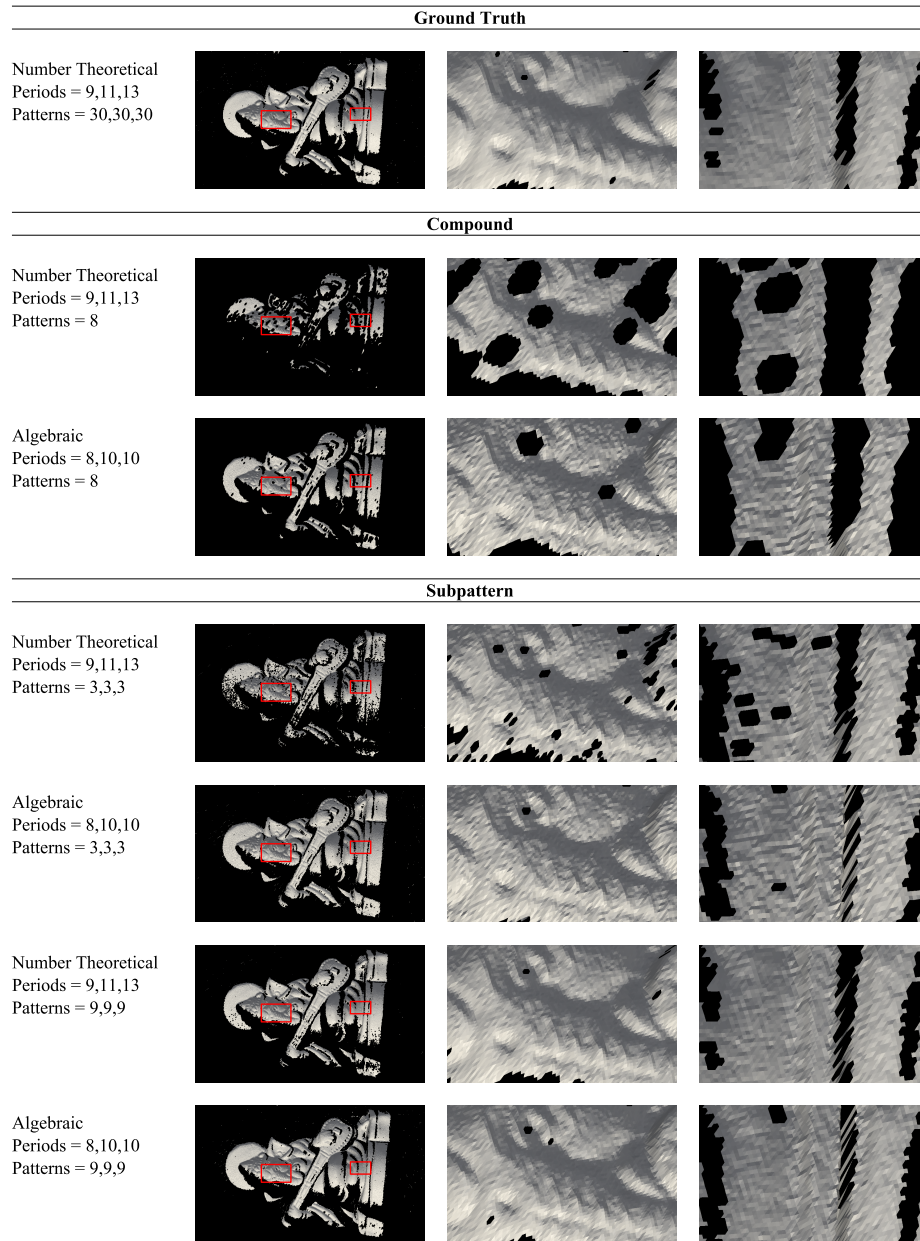Periods = 8,10,10
Patterns = 9,9,9

Figure 11: Qualitative analysis of the surfaces obtained from the light bulb subject.

pattern strategy produces some artifacts (elongated triangles) on the boundary zones. This could be related to the aliasing effect between the bright surface of the bulb and the black background that produces slightly shifted codes which in turn introduce errors in the estimation of the depth of the points. Such artifacts, however, have little influence with respect to the previous quantitative experiments, since they affect only a very limited number of points. Nevertheless, their effect is very significant from a qualitative point of view since such errors stand out very much on the final reconstruction. Finally, the smoothness of the surface increases with the number of patterns, as the effect of the random noise is reduced. Interestingly, with a comparable number of patterns the Algebraic technique seems to give slightly smoother results.

In Fig. 11 we show the results obtained with the Ganesha figurine. It is immediately obvious that both compound methods allow to triangulate much fewer points than the equivalent techniques paired with the subpattern strategy. While the Algebraic coding strategy still gives slightly better results, we can conclude that the subpattern strategy is to be preferred when the object to capture is dark and we can expect a low signal strength. Obviously the subpattern techniques work better with an high number of patterns, with the Algebraic coding being able to reconstruct a few more points. In contrast to the light bulb test, the Algebraic coding schema seems to produce a slightly rougher surface. This could suggest that this method is noisier when dealing with lower signal strength.

6. Conclusions

We have proposed a new fine-to-coarse phase encoding strategy that allows us to achieve high precision and noise robustness with a small number of patterns by applying a coarser angular quantization at the finest level and a finer one at higher levels, and provide a robust estimation approach allowing to recover the projector coordinate with subpixel precision. Further, we have introduced two novel pattern generation strategies that encode multiple phases with a small number of projected patterns and allow to selectively invest more patterns of

signaling band to more critical phases. The proposed approaches have been analyzed and compared with the state-of-the-art, showing that they are at the same time more precise and robust with a high level of noise, not having any significant drawback for low noise levels. In particular, the experiments show that the number-theoretic approach in conjunction with the subpattern generation strategy should be preferred when using a large number of patterns, while the same number-theoretic approach does not perform very well with the compound pattern as proposed in [21] suffers from a relatively high sensitivity to noise. By contrast, the algebraic technique can be used effectively even when projecting a number of patterns close to the Nyquist limit, allowing for a strong reduction in reconstruction time. While the subpattern strategy is still the most robust of the proposed pattern generation strategies, the compound strategy still works relatively well with the algebraic technique in conditions with high signal-to-noise ratio, allowing for good reconstructions with the minimal number of projected patterns.

References

[1] D. Scharstein, R. Szeliski, A taxonomy and evaluation of dense two-frame stereo correspondence algorithm, Int. J. of Comp. Vision 47 (2002) 7–42.

[2] J. Batlle, E. M. Mouaddib, J. Salvi, Recent progress in coded structured light as a technique to solve the correspondence problem: a survey, Pattern Recognition 31 (1998) 963–982.

[3] E. Lilienblum, B. Michaelis, Optical 3D surface reconstruction by a multi-period phase shift method, JCP 2 (2007) 73–83.

[4] M. Takeda, K. Mutoh, Fourier transform profilometry for the automatic measurement of 3-D object shapes, Appl. Opt. 22 (1983) 3977–3982.

[5] J. Salvi, J. Pagès, J. Batlle, Pattern codification strategies in structured light systems, Pattern Recognition 37 (2004) 827–849.

[6] K. L. Boyer, A. C. Kak, Color-encoded structured light for rapid active ranging, IEEE Trans. Pattern Anal. Mach. Intell. 9 (1987) 14–28.

[7] Y. Surrel, Design of algorithms for phase measurements by the use of phase stepping, Appl. Opt. 35 (1996) 51–60.

[8] D. Bergmann, New approach for automatic surface reconstruction with coded light, in: Proceedings of Remote Sensing and Reconstruction for Three-Dimensional Objects and Scenes, volume 2572, pp. 2–9.

[9] G. Wiora, High-resolution measurement of phase-shift amplitude and numeric object phase calculation, in: . A. Y. W. L. J. Latecki, D. M. Mount (Ed.), Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, volume 4117, pp. 289–299.

[10] J. Yi, S. Huang, Modified Fourier transform profilometry for the measurement of 3-D steep shapes, Optics and Lasers in Engineering 27 (1997) 493–505.

[11] K. Hibino, B. F. Oreb, D. I. Farrant, K. G. Larkin, Phase shifting for nonsinusoidal waveforms with phase-shift errors, J. Opt. Soc. Am. A 12 (1995) 761–768.

[12] S. Zhang, S.-T. Yau, Generic nonsinusoidal phase error correction for three-dimensional shape measurement using a digital video projector, Appl. Opt. 46 (2007) 36–43.

[13] B. Pan, Q. Kemao, L. Huang, A. Asundi, Phase error analysis and compensation for nonsinusoidal waveforms in phase-shifting digital fringe projection profilometry, Opt. Lett. 34 (2009) 416–418.

[14] J. Baker, J. F. Chicharo, J. Xi, E. Li., Accuracy limitations introduced by digital projection sources in profilometric optical metrology systems, in: IEEE Conference on Optoelectronic and Microelectronic Materials and Devices, pp. 261–264.

[15] P. S. Huang, Q. J. Hu, F.-P. Chiang, Double three-step phase-shifting algorithm, Appl. Opt. 41 (2002) 4503–4509.

[16] P. L. Wizinowich, Phase shifting interferometry in the presence of vibration: a new algorithm and system, Appl. Opt. 29 (1990) 3271–3279.

[17] Z. Wang, D. A. Nguyen, J. C. Barnes, Some practical considerations in fringe projection profilometry, Optics and Lasers in Engineering 48 (2010) 218–225.

[18] Z. Li, Y. Shi, C. Wang, Y. Wang, Accurate calibration method for a structured light system, Optical Engineering 47 (2008) 053604.

[19] H. Guo, H. He, M. Chen, Gamma correction for digital fringe projection profilometry, Appl. Opt. 43 (2004) 2906–2914.

[20] T. Pribanić, H. Dẑapo, J. Salvi, Efficient and low-cost 3D structured light system based on a modified number-theoretic approach, EURASIP Journal on Advances in Signal Processing 2010 (2010).

[21] A. Albarelli, E. Rodolà, S. R. Bulò, A. Torsello, Fast 3D surface reconstruction by unambiguous compound phase coding, in: IEEE International Workshop on 3D Digital Imaging and Modeling – 3DIM2009, pp. 1670–1677.